(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 February 2005 (10.02.2005)

PCT

(10) International Publication Number
WO 2005/013090 A2

(51) International Patent Classification[7]: G06F

(21) International Application Number: PCT/US2004/024752

(22) International Filing Date: 30 July 2004 (30.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/491,815	1 August 2003 (01.08.2003)	US
60/536,357	14 January 2004 (14.01.2004)	US
60/536,862	15 January 2004 (15.01.2004)	US

(71) Applicant *(for all designated States except US)*: **DNA TWOPOINTO INC.** [US/US]; 1455 Adams Drive, Menlo Park, CA 94025 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GUSTAFSSON, Claes** [US/US]; 2823 Alhambra Drive, Belmont, CA 94002 (US). **GOVINDARAJAN, Sridhar** [IN/US]; 530 Osprey Drive, Redwood City, CA 94065 (US). **MINSHULL, Jeremy, Stephen** [GB/US]; 679 Los Ninos Way, Los Altos, CA 94022 (US).

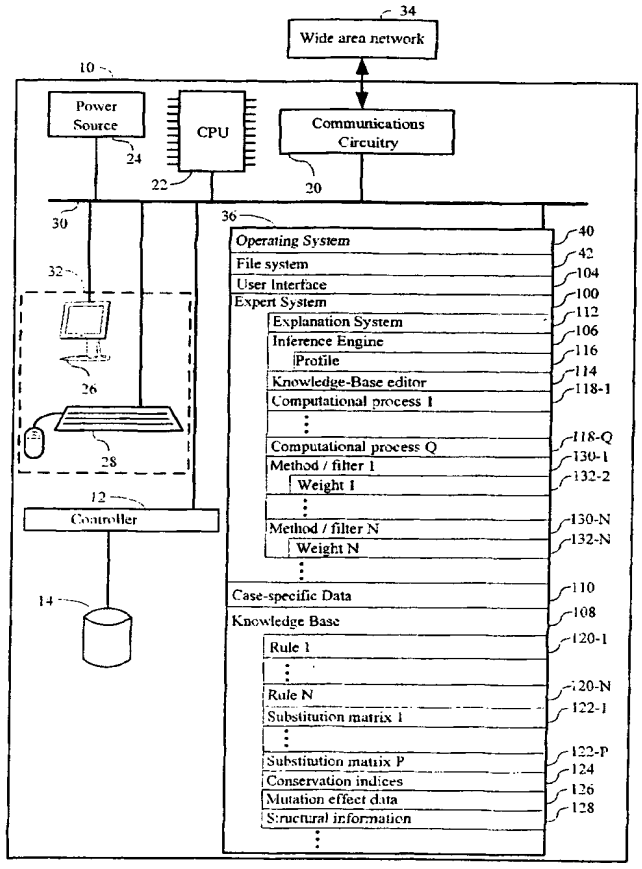
(74) Agents: **LOVEJOY, Brett, A.** et al.; Jones Day, 222 East 41st Street, New York, NY 10017-6702 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH,

[Continued on next page]

(54) Title: SYSTEMS AND METHODS FOR BIOPOLYMER ENGINEERING



(57) Abstract: Methods, computer systems, and computer program products for biopolymer engineering. A variant set for a biopolymer of interest is constructed by identifying, using a plurality of rules, a plurality of positions in the biopolymer of interest and, for each respective position in the plurality of positions, substitutions for the respective position. The plurality of positions and the substitutions for each respective position in the plurality of positions collectively define a biopolymer sequence space. A variant set comprising a plurality of variants of the biopolymer of interest is selected. A property of all or a portion of the variants in the variant set is measured. A sequence-activity relationship is modeled between (i) one or more substitutions at one or more positions of the biopolymer of interest represented by the variant set and (ii) the property measured for all or the portion of the variants in the variant set. The variant set is redefined to comprise variants that include substitutions in the plurality of positions that are selected based on a function of the sequence-activity relationship.

WO 2005/013090 A2

GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.